An Online Image Database of Italian Art. The Database of the Photo Library of the German Art Institute in Florence

Christina Bieber and Werner Schweibenz

Kunsthistorisches Institut in Florenz - Max-Planck-Institut

http://www.khi.fotothek.org

Abstract (EN)

The image database of the photo library of the *German Institute for the History of Art in Florence (Max Planck Institute)* contains some 20,000 high-resolution digital images of Italian works of art with detailed scientific documentation. The paper describes how the project came into being and provides information about the digitization process and the set-up of the information infrastructure. The focus is on specific issues regarding art image databases and the application of the open-source software products ZOPE, PLONE and DIGILIB. According to the institute's guiding principle, which is the promotion of research in art history and the humanities, the image database is managed in compliance with a non-profit policy. For scientific purposes, access to the image database will be available free of charge and copies of the pictures will be provided without publication fees as far as possible.

Keywords (EN): Italian art, image database, digitalisation, photo library, open-source software

Résumé (FR)

La banque d'images de la photothèque du Kunsthistorisches Institut de Florence (Max Plank Institut) relie environ 20.000 photos digitalisées haute résolution d'oeuvres d'art italiennes à une documentation scientifique détaillée. Le rapport décrit le projet, offre des informations quant au processus de digitalisation ainsi que sur la configuration de la chaîne d'informations. L'accent est mis sur les questions propres aux banques de données d'histoire de l'art tout comme sur l'utilisation des produits open source software ZOPE, PLONE et DIGILIB. Conformément à la déontologie de l'Institut, dont le but est de favoriser la recherche en histoire de l'art, l'accès à la banque de données, la consultation et la publication des images seront autant que possible gratuits.

Mots clés (FR): art italien; banque d'images; digitalisation; photothèque: Open-Source-Software

Zusammenfassung (DE)

Die Bilddatenbank der Photothek des Kunsthistorischen Instituts in Florenz (Max Planck Institut) verbindet etwa 20,000 hochaufgelöste digitale Bilder von italienischen Kunstwerken mit einer detaillierten wissenschaftlichen Dokumentation. Der Bericht beschreibt das Projekt, bietet Informationen zum Digitalisierungsprozess und zur Gestaltung der Informationsinfrastruktur. Die Schwerpunkte liegen auf spezifischen Fragen zu kunsthistorischen Bilddatenbanken und der Verwendung der Open-Source-Software-Produkte ZOPE, PLONE and DIGILIB. Entsprechend den Prinzipien des Institutes, das die Förderung der kunsthistorischen Forschung zur Aufgabe hat, wird der Zugang zur Bilddatenbank und die Verwendung der Bilder so weit wie möglich kostenfrei und ohne Publikationsgebühren möglich sein.

Schlüsselwörter (DE): Italienische Kunst, Bilddatenbank, Digitalisierung, Photothek, Open-Source-Software

1. The Photolibrary's Collection and its Digitization Project

The photo library of the *German Institute for Art History in Florence* (*Kunsthistorisches Institut in Florenz* KHI) offers an analogue photo collection of around 600,000 pictures relating to the art of Upper and Central Italy. For some 30,000 photographs the institute holds the copyright which is the legal prerequisite for making them accessible on the Web. Supported by a grant from the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) and the Max Planck Society, in September 2002 the photo library team started a digitazation project in order to make these art photographs digitally accessible on the Internet. Partner in this project was the *Central Institute for the History of Art* (Zentralinstitut für Kunstgeschichte) in Munich, Germany.

The objectives of the project were:

- Choosing the technical specifications and the equipment for digitization
- Selecting the appropriate material
- Creating digital images of the art photographs
- Organizing the data management
- Combining the digital images with the art historical metadata
- Providing online access to the art photographs: For this purpose, a Web presence and image database were created.

2. Choosing the Technical Specifications and Equipment for Digitization

To create digital images of the art photographs, the photo negatives had to be scanned. For conservation reasons, the collection consists almost exclusively of monochrome materials because analogous colour photographs fade relatively quickly in relation to the long-term perspective of a photo library.

First, the technical specifications for the digitization process were established. The goal was to create digital products that can be used for both reproduction and for other purposes in the distant future. Thus a general set of internal standards for future use had to be established, such as options for publication and on-screen display. Therefore different scan resolutions for the different kinds of negative material (glass or film negatives, size, age, or rather physical condition) had to be chosen. The technical specifications for the digitalisation process are presented in table 1.

Table 1 Technical Specifications for the Digitalisation Process

Photo Negative	**Print Media**	**Enlargement**	**Size for Output**	**dpi for Scan**
6x6 cm	Art reproduction 600 dpi	x 4.5	27x27 cm	2700
	Offset reproduction 300 dpi	x 9	54x54 cm	2700
13x18 cm	Art reproduction 600 dpi	x 2.5	32.5x45 cm	1500
	Offset reproduction 300 dpi	x 5	65x90 cm	1500

To implement these technical specifications, high-end technology was required for the entire workflow. The first link in the chain were two high-end scanners. The models chosen were a Flextight 2848 and a Flextight 848, both produced by *Hasselblad Imacon*. The Flextight 2848 can handle negative materials of formats ranging from 35 millimetres to 31x45 centimetres and of different densities such as film and glass. The advantage of the Flextight 2848 is that it has both a virtual drum and a flatbed. So the glass negatives did not have to be copied to film in order to be scanned. The Flextight 848 is a bit smaller and can handle negative materials ranging from 35 millimetres to 12x25 centimetres but no glass negatives. For more technical data on the scanners, see table 2.

Table 2 Technical Data for the Scanners

Scanner Features	Flextight 2848	Flextight 848
Optical Resolution	From 960 to 6,800 dpi	From 80 to 8,000 dpi
Colour Depth	16 bit	16 bit
Density	4.8 Dmax	4.8 Dmax
Scan Speed	100 MB per minute (16 Bit)	100 MB per minute (16/8bit)
Scan Principle	Virtual drum and flatbed	Virtual drum
Film Formats	Transparencies: 35 mm to 29x42 cm (up to 12x17cm virtual drum) Reflectives: up to 31x45 cm (10 mm thick)	Transparencies: 35 mm to 12x25 cm

The scanners are each connected to a *Power Mac G5* with dual processors (64-bit power and high-bandwidth architecture, 2GHz processor speed, 8 GB main memory) for high-speed image processing. Each *G5* is equipped with a 22-inch *LaCie electron22blueIV* monitor.

3. Selecting the Material

For the digitization we had to select the adequate material from the 30,000 photographs for which the KHI holds the copyright. From this material, several photographic archives were selected for digitization. The reasons for this selection were based either on the art historical importance of the collection – for example, the *De Giovanni Archive*, which consists of photographs of the Saint Francis church in Assisi – or on its historic importance – such as the *Lotz Archive*, which contain images depicting the urban development of Florence before the war-inflicted devastations of 1944 or the *Bazzecchi Archive*, which depicts the urban development of Florence before the havoc caused by the great flood of the Arno river in November 1966. On this basis, a set of some 20,000 images was selected for digitization.

4. Creating the Digital Images

The digital images were created in a two-step process. First, a master copy of the photo negative is made with the intent to save as much information as possible. This master copy is used for producing further copies and for long-term preservation. Its size is approximately 100 Megabytes. The technical details are as follows:

- TIFF format uncompressed
- 16-bit colour depth (65,536 colours) per channel

Due to the high colour depth, the file size is relatively large but this had to be accepted in order to gain some leeway for doing the digital work necessary to produce the copies (for example, the working copy).

- Dots per inch (dpi):
  The number of dots per inch is determined, on the one hand, by the informational content of the photo negative and, on the other hand, by the output device and size respectively. In order to achieve the highest possible quality, the degree of dpi was identified separately for each output format and material (cf. Table 3.1 Technical Specifications for the Digitization Process).

A working copy is created from the master scan. This copy is digitally enhanced to allow viewing on screen and printing. Its size is approximately 20 Megabytes. The technical details are as follows:

- TIFF format uncompressed
- 8-bit colour depth (65,536 colours) per channel
- The short side is reduced to 4,000 pixels (approximately 33 cm on a 300 dpi printout)
- 300 dpi

5. Organizing the Data Management

An important concern was how to organize the storage and use of the image data during the project and later on in the image database, we had to develop a system for the file names and coordinate it with the existing numbering system of the photo library's catalogue system. In addition, we had to create a set of directories to structure and access the data on the server. Finally, a sequence of work events was established to coordinate the working process of the photo library and digitization team. This included a final quality control of both the data entered in our art historical meta database and the digitized images.

6. Combining Digital Images with the Art Historical Metadata

The digital images have to be combined with the art historical metadata, which is captured in a meta database system called HiDA (hierarchical data administrator applying the MIDAS rules which are a standard in Germany, cf. Bove et al. 2001). This database system has been used by the KHI for more than ten years and is also used by the consortium in which the KHI is a member. Currently some 15% of the KHI's photo collection has been entered in HiDA and the cataloguing of all new acquisitions in HiDA is to be continued. The metadata in HiDA must be connected to the digital images to provide access for searching and information for interpreting the images. The link between the two data sources is the image file name, which is entered for each HiDA photo document. The HiDA data and the images are stored in the Web database ZOPE, an open-source software that is used for Web publishing.

7 The Image Database

The technical implementation and set-up of the image database were outsourced to a commercial service provider. The prerequisite for the project was to use only open-source software to remain independent of proprietary software. Therefore the Web database ZOPE was chosen for Web publishing (Internet, URL http://zope.org). The texts on the Web site are written in PLINE, the corresponding content management system (Internet, URL

http://plone.org). The following sections describe the most important features of the image database:

- Access to information and images
- The image viewer DIGILIB
- The user interface

7. 1 Access to Information and Images

The access to information and images is provided through three indexes: The artist index, the location index, and the general index.

- The artist index allows searches for artists' names. In addition, it provides an alphabetical overview of all the artists' names in the same form as they can be found in the German standard encyclopaedia Thieme-Becker/Allgemeines Künstlerlexikon.
- The location index consists of an alphabetical list of names of places and locations in Italy. In addition, it allows a search for the different political regions of Italy. Each location link offers a list of all recorded buildings and monuments.
- The general index offers a variety of different search options, for example, date, subject, location, and image content.

The results of a search are displayed in gallery mode. This presentation mode offers a brief description of the works, including information such as artist, date, material, location, and information about the images. The image information consists of two sections. The first section is a gallery view showing thumbnails of the digitised photographs including some basic textual information from HiDA. Each thumbnail is equipped with a link to a larger image and a link to DIGILIB, an image viewer software described in the next section. The second section lists all photographs that are available in the photo library but have not been digitised for different reasons.

7.2 DIGILIB

DIGILIB is an open-source image viewer software developed by the Max Planck Institute for the History of Science, Berlin, the University of Bern, and other partners. Originally designed for digital libraries, DIGILIB lends itself well to image display because most of the functions that were designed for e-books can also be applied to digital images. DIGILIB offers numerous benefits:

- It allows fast transmission of high-resolution digital images (the working copies) because they are broken up into tiles that are transmitted on demand. This allows a reasonable transmission rate for low bandwidth access.
- It allows fast zooms on every detail of the image, so far that you can even see the granularity of the film material.
- It protects high-resolution digital images (the working copies) against copyright violation because the image can be downloaded only in tiles and not as a whole.
- It supports scientific collaboration by providing annotation and marking functions as well as a citation tool for exchange between scholars or students.

These features make DIGILIB a valuable tool for scientific work and collaboration in art history.

FIGURE1: INSERT SCHWEIBENZ_SELECT05-IMAGE1.TIF

Figure 1: An image viewed in DIGILIB

FIGURE2:INSERT SCHWEIBENZ_SELECT05-IMAGE1.TIF

Figure 1: An image viewed by the DIGILIB zoom function

7.3 The User Interface

The user interface of the image database is currently under translation into English and Italian, whereas the content of HiDA must remain in German only. The reasons for this are as follows: First and foremost, HiDA offers no technical support for translation, and, in addition, the database contains too much information to translate without receiving proper funding for now and in the future. Nevertheless, the image database should be usable for scholars and researchers all over the world as the most important search fields such as names of artists and locations follow established conventions and should be understandable for users who do not speak German but have some knowledge of Italian art.

8. Open Access Policy

As a member of the Max Planck Society, the KHI promotes the open-access strategy that has been passed by the Max Planck Society's *Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities* during the 2003 Conference on Open Access to Knowledge in the Sciences and Humanities.

Like the Budapest Open Access Initiative, the ECHO-Charta (European Cultural Heritage Online), or the Budapest Statement on Open Access Publishing before, the German research associations emphasized the vision of a global and accessible representation of knowledge. Therefore the future Web must be sustainable, interactive, and transparent. Content and software tools must be openly accessible and compatible.

Initiatives like these are necessary with regard to the increasingly commercial interests also in the field of cultural heritage and the frequent disappearance of important content from the Web because of financial problems on the part of the maintaining institutions.

We will try to take this into account in our project. But the legal situation in Italy is complicated because of legislation that gives special rights to the owners of works of art. So we first have to establish how far we can follow the open access policy without violating Italian law.

9. Conclusion

The efforts of the whole project were much higher than expected at the beginning. It took a considerable amount of time and effort to sort the archives of the photo library, to design and evaluate the workflow for digitization, and to set up and manage the time-consuming processes of high-end digitalization for materials that were often sensitive and heterogeneous. Equally intensive was the creation of the new database and the Web presence. In particular, the information design required some thought: How to set up the structure of the Web site and how to present the necessary content. Nevertheless, from our perspective it was worth the effort as we can now present the collection of the KHI photo library on the Web.

References

Bove, J. & Heusinger, L. & Kailus, A. (2001). Marburger Informations-, Dokumentations- und Administrations-System (MIDAS). München: Saur. (Title in english: Manual of the MIDAS Rules)

Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities. The Max Planck Society's Conference on Open Access to Knowledge in the Sciences and Humanities, 20 - 22 Oct 2003, Berlin. Available <http://www.zim.mpg.de/openaccess-berlin/berlindeclaration.html>.

Authors

Christina Bieber, Werner Schweibenz
Fototeca
Kunsthistorisches Institut in Florenz / *German Institute for Art History in Florence*
Via G. Giusti 44
I-50121 Firenze, Italy
phone: 0039-055-24911-39
fax: 0039-055-24911-76
bieber@khi.fit.it, schweibenz@khi.fi.it

Christina Bieber studied documentation science at the university of applied sciences at Cologne and information science with art history and economics as minors at the University of Saarland at Saarbrücken, Germany. Since September 2002 she works for the German Art Institute' photo library as information specialist. She acted as the technical lead of the documentation project described in the article.

Werner Schweibenz studied information science with art history and legal informatics as minors at the University of Saarland at Saarbrücken, Germany. Since April 2004 he is working as information specialist for the photo library of the German Art Institute in Florence. In addition he works on his ph.d. project on virtual museums and image databases.